CrossMark

On a measure with maximal entropy for a suspension flow over a countable alphabet Markov shift

Boris Gurevich[1,2]

Abstract Sufficient conditions on a shift invariant probability measure on the base of a suspension flow $\{S_t\}$ over a countable alphabet Markov shift are stated under which the corresponding $\{S_t\}$-invariant measure is a measure with maximal entropy.

Keywords Topological Markov shift · Suspension flow · Topological entropy · Measure with maximal entropy

Mathematics Subject Classification 37A10 · 37A35 · 37B10 · 60J10

1 Introduction

The present paper is motivated by the study in [1,2] of a measure with maximal entropy for the Teihmüller flow on the moduli space of Abelian differentials. The following result was obtained there: a measure with maximal entropy for this flow exists, is unique, and moreover, it is the well-known invariant measure constructed by H. Masur and W. Veech in the 80th. In several steps the problem reduces to the investigation of a measure with maximal entropy for a suspension flow over a countable alphabet Markov shift. Our argument can be briefly presented as follows. Veech [12] found the

✉ Boris Gurevich
bmgbmg2@gmail.com

1 Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Leninskie Gory, GSP-1, 119991 Moscow, Russia

2 Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoy Karetny per. 19, Building 1, 127051 Moscow, Russia

entropy of the Teihmüller flow with respect to the Masur–Veech measure and expressed it in terms of the genus g of the corresponding Riemann surface and the number n of punctures (in the moduli space context, the surface itself is not fixed, but g and n are). For the measure μ obtained by transferring the Masur–Veech measure to the phase space of the suspension flow in question (more exactly to its base), we discovered some properties in terms of this flow alone, independently of the origin of μ. Using the existence of an invariant measure with these properties, we evaluated the topological entropy (the maximum of the measure theoretic entropies) of the suspension flow, which turned out to coincide with the Veech entropy. This immediately implied our main result.

Now the following natural question arises: if Veech's entropy result [12] have not been known, would it be possible to deduce it directly from the properties of the measure, without resorting to the Teihmüller flow itself? In other words, taking into account that our evaluation of the topological entropy is valid not only for the specific suspension flow related to the Teihmüller flow, but for every suspension flow that has an invariant measure with the above properties, is it true that the entropy of this flow with respect to such a measure coincides with its topological entropy? It could seem that the whole information contained in the assumptions on μ was already used in [2]. In fact this is not true. In the present paper we are based on the same assumptions as in [2], but exploit them in a different way (see the proof of Theorem 2.10 below).

2 Definitions and results

We start from some more or less standard definitions.

Definition 2.1 A one-to-one bi-measurable transformation T of a measurable space $(X, \mathcal{B})$ onto itself is called an *automorphism*, a group of automorphisms T^t, $t \in \mathbb{Z}$, is called a *cascade* (a discrete time dynamical system); a continuous parameter group of automorphisms T^t, $t \in \mathbb{R}$, is called a *flow* (a continuous time dynamical system). Two dynamical systems, $(X_1, \mathcal{B}_1, \{T_1^t\})$ and $(X_2, \mathcal{B}_2, \{T_2^t\})$, are called *isomorphic* if there is a one-to-one epimorphic bi-measurable map $\Phi \colon X_1 \to X_2$ such that $T_2^t \circ \Phi = \Phi \circ T_1^t$ for all t. (For the discrete time dynamical systems, it is sufficient to check the last equation for $t = 1$.)

We often have to deal with a dynamical system $(X, \mathcal{B}, \{T^t\})$ endowed with a $\{T^t\}$-invariant probability measure μ on X. We will call the quadruple $(X, \mathcal{B}, \{T^t\}, \mu)$ a dynamical system with invariant measure.

Definition 2.2 Two dynamical systems with invariant measure, $(X_1, \mathcal{B}_1, \{T_1^t\}, \mu_1)$ and $(X_2, \mathcal{B}_2, \{T_2^t\}, \mu_2)$, are called *isomorphic* if there are sets $X_1' \in \mathcal{B}_1$, $X_2' \in \mathcal{B}_2$ invariant with respect to all T_1^t and all T_2^t, respectively, such that $\mu_1(X_1') = \mu_2(X_2') = 1$ and the restrictions $(X_1', \mathcal{B}_1|_{X_1'}, \{T_1^t\}|_{X_1'})$ and $(X_2', \mathcal{B}_2|_{X_2'}, \{T_2^t\}|_{X_2'})$ are isomorphic in the sense of Definition 2.1.

In what follows we will usually drop $\mathcal{B}$ from the notation of a dynamical system.

A *suspension flow* (or *special flow*) over an automorphism T is defined as follows (see, e.g., [4,8]).

Definition 2.3 Let T be an automorphism of a measurable space $(X, \mathcal{B})$, and $f : X \to (c, \infty)$, $c > 0$, be a Borel measurable function. We set

$$X_f = \{\tilde{x} = (x, u) : x \in X, \ 0 \leq u < f(x)\}.$$

For $t \geq 0$ and every $\tilde{x} = (x, u) \in X_f$, we define a map S_t by

$$
S_t \tilde{x} = \begin{cases} (x, u + t) & \text{if } u + t < f(x), \\ \left(T^n x, \ u + t - \displaystyle\sum_{i=0}^{n-1} f(T^i x) \right) & \text{otherwise,} \end{cases}
$$

where n is such that

$$\sum_{i=0}^{n-1} f(T^i x) \leq u + t < \sum_{i=0}^{n} f(T^i x).$$

For $t < 0$ we set $S_t = (S_{-t})^{-1}$. For the suspension flow thus obtained we will use the notation (T, f). The space X and the function f are often called the base of the suspension flow and the roof function, respectively.

The notion of a suspension flow is an abstract version of the Poincare transversal surface idea. Suspension constructions (flows and cascades) over countable alphabet Markov shifts turn out to be a powerful tool in the study of some classes of smooth dynamical systems on manifolds [1,2,6,10].

For an automorphism T of X and for $x \in X, C \subset X$, we set

$$\tau(x, C) = \inf \{n > 0 : T^n x \in C\}, \qquad x \in X, \tag{1}$$

and for the suspension flow $\{S_t\} = (T, f)$, we set

$$\tilde{\tau}(x, C) = \inf \{t > 0 : S_t(x, 0) \subset C \times \{0\}\};$$

$\tau(x, C)$ is clearly the first approach time of the set C by the point x under the action of the cascade $\{T^n\}$, while $\tilde{\tau}(x, C)$ can be a little inaccurately treated as the first approach time of the set C by the point x under the action of the flow (T, f) (we put $\tau(x, C) = \infty$ if x will never approach C).

We are mainly interested in suspension flows over topological (or symbolic) Markov shifts.

Definition 2.4 Let $G = (V, E)$ be a directed graph with vertex set V and edge set $E \subset V \times V$. One defines the two-sided infinite path space $X(G) = X$ of G by

$$X = \{x = (x_i, \ i \in \mathbb{Z}) : x_i \in V, \ (x_i, x_{i+1}) \in E\}$$

and introduces the topology of pointwise convergence in $X(G)$ (every $x \in X$ is a function from $\mathbb{Z}$ to V) and the (left) shift transformation T by

$$(Tx)_i = x_{i+1}, \qquad i \in \mathbb{Z}, \quad x \in X.$$

This T, which is clearly a homeomorphism of X, is called the *topological* (*or symbolic*) *Markov shift* with alphabet V. If $G = (V, E)$ is a complete graph, i.e., $V = G \times G$, the shift T on $X(G)$ is called a *full shift* or a *topological* (*symbolic*) *Bernoulli shift*.

In what follows we assume that V is countable and G is connected as a directed graph. Let

$$\operatorname{var}_n^+(f) = \sup \left\{ |f(x) - f(x')| : x, x' \in X, \ x_i = x'_i \text{ for } 0 \le i \le n \right\}.$$

We impose the following conditions on f: $f(x)$ does not depend on x_i with $i < 0$ and

$$\sum_{n=1}^{\infty} \operatorname{var}_n^+(f) < \infty.$$

Remark 2.5 While the former condition can be essentially relaxed, it enables us to use results from [2], where this condition is present. In contrast, the latter condition, which is also present in [2], is significant.

Denote the family of all T-invariant Borel probability measures on X by $\mathfrak{I}(T)$ and set

$$\mathfrak{I}_f(T) = \{\mu \in \mathfrak{I}(T) : \mu(f) < \infty\}$$

(here and below we write $\mu(f)$ instead of $\int_X f \, d\mu$),

$$\mathfrak{I}_f^+(T) = \left\{ \mu \in \mathfrak{I}_f(T) : \mu(C) > 0 \text{ for each open } C \subset X \right\}.$$

If $\mu \in \mathfrak{I}_f(T)$, it generates, in a canonical way, a probability measure μ_f on X_f that is invariant w.r.t. the suspension flow $\{S_t\} = (T, f)$.

For $\mu \in \mathfrak{I}_f(T)$, denote the entropy of T with respect to μ by $h_\mu(T)$ and the entropy of the flow (T, f) with respect to the measure μ_f by $h_{\mu_f}(T, f)$.

Definition 2.6 Let

$$h_{\text{top}}(T, f) = \sup_{\mu \in \mathfrak{I}_f(T)} h_{\mu_f}(T, f).$$

We say that $h_{\text{top}}(T, f)$ is the *topological entropy* of the flow (T, f) and that every measure $\mu \in \mathfrak{I}_f(T)$ with

$$h_{\mu_f}(T, f) = h_{\text{top}}(T, f)$$

is a *maximal measure* for (T, f).

Remark 2.7 The justification for the above definition of the topological entropy of (T, f) is that, in the case of a finite alphabet Markov shift, this quantity would really be the topological entropy of this suspension flow.

Remark 2.8 It is known [2] that in the situation we are going to study, a maximal measure is unique, but even in quite simple cases it does not necessarily exist.

In order to state the main result of this paper we have to give the following definition.

Definition 2.9 Given a graph $G = (V, E)$ as above, we call a sequence $w = (v_0, v_1, \ldots, v_{n-1})$ with $v_i \in V$, $(v_i, v_{i+1}) \in E$ a *path* (or a *word*) of length $|w| = n$ in G. We call such w a *simple word* if $(v_0, v_1, \ldots, v_k) \neq (v_{n-k}, v_{n-k+1}, \ldots, v_n)$ unless $k = n$. We denote the set of all words in G by $W(G)$.

If $w = (v_0, v_1, \ldots, v_k)$ and $w' = (v'_0, v'_1, \ldots, v'_l)$ are two words in G such that $v_k = v'_0$, then their concatenation $ww' = (v_0, v_1, \ldots, v_k, v'_1, \ldots, v'_l)$ is also a word in G. We write $w' \subset w$ if $v_{r+i} = v'_i$ for some $r \geq 0$ and $i = 0, 1, \ldots, l$. For every $w \in W(G)$, we denote the cylinder $C_w = \{x \in X : (x_0, \ldots, x_{|w|-1}) = w\}$ by C_w.
 The next theorem is the main result of the paper. To state it we have to introduce the following four conditions on a measure $\mu \in \mathfrak{I}_f(T)$:

(a) $\mu \in \mathfrak{I}_f^+(T)$, i.e., $\mu(C) > 0$ for all open nonempty $C \subset X$.
(b) μ is ergodic, i.e., the transformation T is ergodic with respect to the measure μ.
(c) $H_\mu(\xi) < \infty$, where ξ is the partition of X into one-dimensional cylinders C_v, $v \in V$, and $H_\mu(\xi)$ is the entropy of this partition with respect to μ.
(d) There exist a constant $s \geq 0$, a function $\kappa : \mathbb{Z}_+ \to \mathbb{R}_+$, and an arbitrary long simple word $w \in W(G)$ such that $\lim_{r \to \infty} \kappa(r) = 0$ and, for all $\widehat{w} \in W(G)$ satisfying the conditions $w\widehat{w}w \in W(G)$, $\widehat{w} \not\subset w$ and μ-almost all $x \in C_{ww\widehat{w}}$, the inequality

$$\left| \frac{\mu(C_{ww'w})}{\mu(C_w)} - e^{-s\widetilde{\tau}(x, C_w)} \right| \leq \kappa(|w|) e^{-s\widetilde{\tau}(x, C_w)}$$

holds. Each word $\widehat{w}$ satisfying two last conditions will be called w-*admissible*, we denote the set of such words by $\widehat{A}_w$.

Theorem 2.10 *Let (X, T) be an infinite alphabet Markov shift determined by a connected graph $G = (V, E)$ and f be a function with summable variations defined on $X(G)$ such that $f(x)$ depends on $x_0, x_1, \ldots$ only and $f(x) \geq c > 0$. Assume that a measure $\overline{\mu} \in \mathfrak{I}_f(T)$ satisfies conditions* (a)–(d). *Then $\overline{\mu}$ is a maximal measure for the suspension flow (T, f).*

Remark 2.11 This theorem is close to [2, Theorem 2.2]. In particular, the conclusions of these theorems coincide. But their assumptions are different: instead of the equality $s = h_{\overline{\mu}_f}(T, f)$, we assume that $H_{\overline{\mu}}(\xi) < \infty$. The second assumption seems to be much less restrictive and purely technical. It is quite possible that it can be rejected. As to the first assumption, we derive it from the other assumptions of Theorem 2.10. Moreover, there is good reason to believe that the sufficient conditions on the maximal measure stated in the theorem are close to necessary ones.
 Theorem 2.10 will be proved in Sect 4. In Sect. 3 we state and in part prove some auxiliary facts.

3 Main lemmas

Lemma 3.1 *Let (T, f) be a suspension flow over an infinite alphabet topological Bernoulli shift (X, T), where $f(x)$ depends only on x_0, i.e., there exists a function $f_0 \colon V \to [c, \infty), c > 0$, such that $f(x) = f_0(x_0), x \in X$. Let*

$$F_{f_0}(\beta) = \sum_{v \in V} \exp(-\beta f_0(v)), \qquad \beta > 0.$$

Then

(i) *$h_{\text{top}}(T, f)$ is the unique solution to the equation $F_{f_0}(\beta) = 1$ if this equation is solvable, otherwise*

$$h_{\text{top}}(T, f) = \sup\{\beta \geq 0 : F_{f_0}(\beta) = \infty\}.$$

(ii) *(T, f) has a maximal measure $\mu_{\max}$ if and only if there exists a number $\beta_0 > 0$ such that*

$$F_{f_0}(\beta_0) = 1, \qquad |F'_{f_0}(\beta_0)| < \infty.$$

This measure is unique and is the Bernoulli measure such that $\mu_{\max}(C_v) = \exp(-\beta_0 f_0(v)), v \in V$.

Statement (i) is [2, Lemma 2.7] (see also [5]), statement (ii) can be derived from (i) or from [11, Theorem 3].

Lemma 3.2 *Let (T, f) be a suspension flow as in Lemma 3.1 and $h = h_{\text{top}}(T, f)$. Then, for the function $g(x) = f(x) - h^{-1} \ln F_{f_0}(h)$,*

$$F_g(h) = 1, \qquad h_{\text{top}}(T, g) = h \tag{2}$$

(g_0 corresponds to g in the same way as f_0 corresponds to f).

Proof From Lemma 3.1 (i) it follows that either $F_{f_0}(h) = 1$ or $F_{f_0}(h) < 1$. Clearly in both cases (2) holds. □

Lemma 3.3 *Let (T, f) be a suspension flow over an infinite alphabet topological Bernoulli shift (X, T) and $\inf_{x \in X} f(x) \geq c > 0$. Assume that for every $v \in V$, every pair $x, x' \in C_v$, and some $\delta \in [0, c/2]$, we have*

$$|f(x) - f(x')| \leq \delta.$$

Then there exists a function $g \colon X \to [0, \infty)$ such that

(α) *$g(x) = g_0(x_0)$ for some function $g_0 \colon V \to [0, \infty)$;*
(β) *$|g(x) - f(x)| \leq \delta, x \in X$;*
(γ) *$h_{\text{top}}(T, f) = h_{\text{top}}(T, g)$.*

Proof Let

$$f^+(x) = \sup_{\substack{x' \in X \\ x'_0 = x_0}} f(x'), \qquad f^-(x) = \inf_{\substack{x' \in X \\ x'_0 = x_0}} f(x'), \qquad x \in X.$$

Clearly

$$f^-(x) \le f(x) \le f^+(x), \qquad 0 \le f^+(x) - f^-(x) \le \delta.$$

Hence

$$h_{\text{top}}(T, f^+) \le h_{\text{top}}(T, f) \le h_{\text{top}}(T, f^+ - \delta). \tag{3}$$

Moreover,

$$|f^+(x) - t - f(x)| \le \delta, \qquad x \in X, \quad t \in [0, \delta]. \tag{4}$$

We want to show that the function $t \mapsto h_{\text{top}}(T, f^+ - t)$ is continuous in $t \in [0, c/2]$. For $t' \in [0, c/4]$, we obviously have

$$
\begin{aligned}
0 &\le h_{\text{top}}(T, f^+ - t - t') - h_{\text{top}}(T, f^+ - t) \\
&= \sup_{\mu \in \mathcal{J}(T)} \frac{h_\mu(T)}{\mu(f^+) - t - t'} - \sup_{\mu \in \mathcal{J}(X,T)} \frac{h_\mu(T)}{\mu(f^+) - t} \\
&\le \sup_{\mu \in \mathcal{J}(T)} h_\mu(T) \left[\frac{1}{\mu(f^+) - t - t'} - \frac{1}{\mu(f^+) - t} \right] \\
&\le t' \sup_{\mu \in \mathcal{J}(T)} \frac{h_\mu(T)}{(\mu(f^+) - t - t')(\mu(f^+) - t)} \\
&\le t' \sup_{\mu \in \mathcal{J}(T)} h_\mu(T) \frac{8}{3c\mu(f)} = t' \frac{8}{3c} h_{\text{top}}(T, f).
\end{aligned}
\tag{5}
$$

In the same way and for the same t, t', we obtain

$$0 \le h_{\text{top}}(T, f^+ - t) - h_{\text{top}}(T, f^+ - t + t') \le \frac{4t'}{c} h_{\text{top}}(T, f). \tag{6}$$

From (5) and (6) the continuity required follows.

Now, using (3), (4) and the assumption $\delta \le c/2$, we can find a point $t_0 \in [0, \delta]$ such that

$$h_{\text{top}}(T, f^+ - t_0) = h_{\text{top}}(T, f), \qquad \sup_{x \in X} |f^+(x) - t_0 - f(x)| \le \delta.$$

It is evident that the function $g(x) = f^+(x) - t_0$, $x \in X$, satisfies conditions (α)–(γ) above.   □

The following lemma (see [6]) will be used in the end of the next section.

Lemma 3.4 *Let* $\{\mu_n, n \geq 0\}$ *be a sequence of* T-*invariant probability measures on* $X = X(G)$, *see Introduction. Assume that there exist two sequences,* r_n *and* $\alpha_n, r_n \in \mathbb{N}, \alpha_n \in \mathbb{R}_+, n \in \mathbb{N}$, *and a constant* $h \geq 0$ *such that*

$$\lim_{n \to \infty} r_n = \infty, \qquad \lim_{n \to \infty} \alpha_n = 0, \qquad \lim_{n \to \infty} h_{\mu_n}(T) = h,$$

$$|\mu_0(C) - \mu_n(C)| \leq \alpha_n \mu_n(C)$$

for all cylinders $C \subset X$ *of the form* $C = C_{\widetilde{w}}$, *where* $\widetilde{w}$ *is a word in* G *of length* $|\widetilde{w}| = r_n$. *Then* $h_{\mu_0}(T) \geq h$.

4 Proof of Theorem 2.10

We divide the proof into several steps.

1. For a simple word $w \in W(G)$, we set

$$C_w^\infty = \left\{ x \in X : \sum_{i=0}^\infty \mathbf{1}_{C_w}(T^{-i}x) = \sum_{i=0}^\infty \mathbf{1}_{C_w}(T^i x) = \infty \right\}, \qquad (7)$$

i.e., C_w^∞ is the set of the points $x \in X$ that approaches C_w infinitely many times under the action of both T and T^{-1}. This set is clearly measurable and T-invariant. So one can consider the restriction $T|_{C_w^\infty}$ and the suspension flow $(T|_{C_w^\infty}, f|_{C_w^\infty})$.

Let $C_w' = C_w^\infty \cap C_w$ and $T_{C_w'}$ be the transformation induced by T on C_w' (the first return map).

The following statements are rather simple and essentially known (see Lemma 2.8 and the proof of Theorem 2.2 in [2]).

Proposition 4.1 *The dynamical system* $(C_w', T_{C_w'})$ *is isomorphic to the Bernoulli shift with alphabet* A_w, *consisting of all words* $w\widehat{w}w \in W(G)$, *where* $w \not\subset \widehat{w}$.

Proof Taking into account the simplicity of w, one sees that every point $x \in C_w'$ is a sequence $(x_i)_{i \in \mathbb{Z}}$ such that, for each $k \in \mathbb{Z}$, there exists $i_k = i_k(x) \in \mathbb{Z}$ with the following properties: 1) $i_0 = 0$, 2) $i_{k+1} \geq i_k + |w|$, 3) $(x_{i_k}, \ldots, x_{i_k+|w|-1}) = w$, 4) $(x_{i_k+|w|-1}, \ldots, x_{i_{k+1}}) = \widehat{w}$. It is clear, that $T_{C_w'}x$ is the left shift of x by $|w| + |\widehat{w}_k| - 2 = i_{k+1} - i_k - 1$ steps. We assign the sequence $\Phi_w x = (w\widehat{w}_k w)_{k \in \mathbb{Z}}$ to x and observe that Φ_w is a continuous one-to-one map from C_w' to $(A_w)^{\mathbb{Z}}$. Moreover, $\Phi_w T_{C_w'}x = T^{|w|+|\widehat{w}_k|-2}\Phi_w x$. $\qquad\square$

Let

$$f_w(x) = \sum_{i=0}^{\tau(x,C_w)-1} f(T^i x), \qquad x \in C_w'.$$

(See (1) for the definition of $\tau(x, C_w)$.)

Proposition 4.2 *If $w\widehat{w}w \in A_w$ and $x, x' \in C_{w\widehat{w}w}$, then*

$$|f_w(x) - f_w(x')| \leq \sum_{n=|w|}^{\infty} \text{var}_n^+(f). \tag{8}$$

Proposition 4.3 *The suspension flows $(T|_{C_w^\infty}, f|_{C_w^\infty})$ and $(T_{C_w'}, f_w)$ are isomorphic.*

Proof By Definition 2.3 the flow $(T|_{C_w^\infty}, f|_{C_w^\infty})$ acts on the space $(C_w^\infty)_f$ that consists of the pairs (x, u), where $x \in C_w^\infty, 0 \leq u < f(x)$. For every such a pair, we set

$$m(x) = \min\{i \geq 0 : T^{-i}x \in C_w\}, \qquad x' = T^{-m(x)}x,$$

$$u' = \begin{cases} u & \text{if } m(x) = 0, \\ u + \sum_{k=0}^{m(x)-1} f(T^k x) & \text{if } m(x) \geq 1. \end{cases}$$

It is easy to see that $\widetilde{\Phi}_w \colon (x, u) \mapsto (x', u')$ is a bi-measurable one-to-one map of $(C_w^\infty)_f$ to $(C_w')_f$ that takes $(T|_{C_w^\infty}, f|_{C_w^\infty})$ to $(T_{C_w'}, f_w)$. $\qquad\square$

Corollary 4.4 $h_{\text{top}}(T|_{C_w^\infty}, f|_{C_w^\infty}) = h_{\text{top}}(T_{C_w'}, f_w) = h_{\text{top}}(T, f)$.

Proof The first equality is obvious, the second one follows from the fact that the supremum in Definition 2.6 can be taken over $\mathcal{I}_f^+(T)$, see, e.g., [2, Lemma 2.9]. $\quad\square$

Denote the sequence space $(A_w)^{\mathbb{Z}}$ by Y_w, its elements by $y = (y_i)_{i \in \mathbb{Z}}$, and the one step left shift on Y_w by σ. Proposition 4.1 implies that the suspension flow $(T_{C_w'}, f_w)$ is isomorphic to the suspension flow (σ, φ^w), where

$$\varphi^w(y) = f_w(\Phi^{-1}y), \qquad y \in Y_w. \tag{9}$$

Denote the right hand side of (8) by δ_w. By (8) and (9)

$$|\varphi^w(y) - \varphi^w(y')| \leq \delta_w$$

for all $y, y' \in Y_w$ with $y_0 = y_0'$.

2. Let $\mathcal{W}(G)$ be the set of simple words for which condition (d) on the measure $\overline{\mu}$ (see Introduction) is satisfied. For each $w \in \mathcal{W}(G)$, we will introduce two probability measures, μ^w and ν^w, on Y_w. To define μ^w we note that $\overline{\mu}(C_w) > 0$ (see condition (a) on $\overline{\mu}$) and hence $\overline{\mu}(C_w') > 0$. Now denote $\mu^w = \Phi_w \circ \overline{\mu}_w'$ where $\overline{\mu}_w'$ is the normalized restriction of $\overline{\mu}$ to C_w'.

To define ν^w, we first observe that by the definition of the functions f_w and φ^w the condition (d) on $\overline{\mu}$ can be rewritten in the form

$$\left|\mu^w(C_a) - e^{-s\varphi^w(y)}\right| \leq \kappa(|w|)e^{-s\varphi^w(y)} \tag{10}$$

for all $a \in A_w$ and μ^w-almost all $y \in C_a$.

Now choose $w \in \mathcal{W}(G)$ such that

$$\delta_w < \frac{c}{2} \leq \frac{1}{2} \inf_{y \in Y_w} \varphi^w(y)$$

(the latter inequality follows from the definition of φ^w). Then we can apply Lemma 3.3 to find a function ψ^w on Y_w such that:

(α_w) $\psi^w(y) = \psi_0^w(y_0)$ for some function $\psi_0^w : A_w \to [0, \infty)$;
(β_w) $|\psi^w(y) - \varphi^w(y)| \leq \delta_w$, $y \in Y_w$;
(γ_w) $h_{\text{top}}(\sigma, \varphi^w) = h_{\text{top}}(\sigma, \psi^w)$.

The next proposition is a statement of [2, Theorem 2.2 (i)], one can observe that only conditions (a), (b) and (d) on the measure $\overline{\mu}$ are used in its proof.

Proposition 4.5 *Under the conditions of Theorem* 2.10 *the equation* $s = h_{\text{top}}(T, f)$ *holds.*

We now perturb a little the function ψ^w to add one useful property to (α_w)–(γ_w).

Proposition 4.6 *If* $|w|$ *is large enough, then there exists a number* r_w *such that*

$$0 \leq r_w \leq -\frac{1}{s} \log \frac{e^{-s\delta_w}}{1 + \kappa(w)}, \tag{11}$$

$$h_{\text{top}}(\sigma, \varphi^w) = h_{\text{top}}(\sigma, \psi^w - r_w), \tag{12}$$

$$\sum_{a \in A_w} \exp[-s\psi_0^w(a) - r_w] = 1. \tag{13}$$

Proof For every $a \in A_w$, we can choose a point $y_a \in C_a$ such that

$$\left| \mu^w(C_a) - e^{-s\varphi^w(y_a)} \right| \leq \kappa(|w|) e^{-s\varphi^w(y_a)}$$

(see (10)) and hence

$$e^{-s\varphi^w(y_a)} \geq \frac{\mu^w(C_a)}{1 + \kappa(|w|)}. \tag{14}$$

By (α_w), (β_w) and (14)

$$\exp\left[-s\psi_0^w(a)\right] = \exp\left[-s\big(\varphi^w(y_a) + \psi_0^w(a) - \varphi^w(y_a)\big)\right]$$
$$\geq e^{-s\varphi^w(y_a)} e^{s\delta_w} \geq \frac{e^{s\delta_w} \mu_w(C_a)}{1 + \kappa(|w|)}. \tag{15}$$

Summing up (15) over a and taking into account Lemmas 3.1, 3.2, Proposition 4.4 and the fact that μ^w is a probability measure, we obtain

$$\frac{e^{-s\delta_w}}{1 + \kappa(|w|)} \leq \sum_{a \in A_w} e^{-s\psi_0^w(a)} \leq 1.$$

If we now set

$$r_w = -\frac{1}{s} \log \sum_{a \in A_w} e^{-s\psi_0^w(a)},$$

then (11)–(13) will clearly hold. □

Corollary 4.7 *For any $\varepsilon > 0$, there exists a word $w = w_\varepsilon \in \mathcal{W}(G)$ so long that one can find a function $\psi^w : Y_w \to (0, \infty)$ with the following properties:*

$$
\begin{aligned}
&\psi^w(y) = \psi_0^w(y_0) &&\text{for some function } \psi_0^w : A_w \to [0, \infty); \\
&|\psi^w(y) - \varphi^w(y)| \leq \varepsilon &&\text{for all } y \in Y_w; &&(16)\\
&h_{\text{top}}(\sigma, \varphi^w) = h_{\text{top}}(\sigma, \psi^w); &&&&(17)\\
&\sum_{a \in A_w} \exp[-s\psi_0^w(a)] = 1.
\end{aligned}
$$

This follows immediately from Proposition 4.6 due to the properties of the function ψ^w there (see (α_w)–(γ_w) above) and the fact that

$$\lim_{|w|\to\infty} \delta_w = \lim_{|w|\to\infty} \kappa(|w|) = 0$$

(for the new function we use the same notation ψ^w).

From now on we assume that ε and $w = w_\varepsilon$ as in Corollary 4.7 are fixed, unless otherwise stated.

We define ν^w to be the Bernoulli measure on Y_w with

$$\nu^w(C_a) = \exp\left[-s\psi_0^w(a)\right], \qquad a \in A_w. \qquad (18)$$

Let us show that the measures μ^w and ν^w are in some sense close to each other. From (10) and the definition of ν^w, for every $a \in A_w$, one obtains

$$
\begin{aligned}
|\mu^w(C_a) - \nu^w(C_a)| &\leq |\mu^w(C_a) - e^{-s\varphi^w(y_a)}| + |e^{-s\varphi^w(y_a)} - e^{-s\psi_0^w(a)}| \\
&\leq e^{-s\varphi^w(y_a)}\left(\kappa(|w|) + |1 - e^{-s(\psi_0^w(a)-\varphi^w(y_a))}|\right) \qquad (19)\\
&\leq \nu^w(C_a)e^{-s(\varphi^w(y_a)-\psi_0^w(a))}\left(\kappa(|w|) + |1 - e^{s(\varphi^w(y_a)-\psi_0^w(a))}|\right).
\end{aligned}
$$

Taking w still longer if necessary, from (19), (16) and the condition on $\kappa(|w|)$ we obtain

$$|\nu^w(C_a) - \mu^w(C_a)| \leq \varepsilon\nu^w(C_a), \qquad a \in A_w. \qquad (20)$$

From (16) and (20) we see that $\nu^w(\psi^w) < \infty$. Hence Lemma 3.1 and Proposition 4.5 combined with (18) show that ν^w is the maximal measure for the suspension flow (σ, ψ^w).

3. We now carry over the measures ν^w and μ^w to X with the help of the map Φ_w^{-1} (see the proof of Proposition 4.1). So we set

$$\nu'_w = \Phi_w^{-1} \circ \nu^w.$$

We also have $\overline{\mu}_w = \Phi_w^{-1} \circ \mu^w$ (see step 2 of this proof).

From the definition of Φ_w it follows that ν'_w is a probability measure on C'_w, and (20) can be rewritten, for every $\widehat{w} \in \widehat{A}_w$, as

$$\left| \nu'_w\left(C'_w \cap C_{w\widehat{w}w}\right) - \overline{\mu}'_w(C_{w\widehat{w}w}) \right| \le \varepsilon \nu'_w\left(C'_w \cap C_{w\widehat{w}w}\right). \tag{21}$$

The next goal is to extend ν'_w to the whole X. We first observe that

$$C'_w = \bigsqcup_{\widehat{w} \in \widehat{A}_w} \left(C'_w \cap C_{w\widehat{w}w}\right). \tag{22}$$

Moreover, the sets $T^i\left(C'_w \cap C_{w\widehat{w}w}\right)$ do not intersect each other when $\widehat{w} \in \widehat{A}_w$ and $i = 0, 1, \ldots, |w| + |\widehat{w}| - 3$. It is easy to check that

$$C_w^\infty = \bigsqcup_{\widehat{w} \in \widehat{A}_w} \bigsqcup_{i=0}^{|w|+|\widehat{w}|-3} T^i\left(C'_w \cap C_{w\widehat{w}w}\right),$$

where C_w^∞ is defined in (7).

Taking into account (22), we introduce a measure ν_w^∞ on C_w^∞ by

$$\nu_w^\infty|_{T^i(C'_w \cap C_{w\widehat{w}w})} = T^i \circ \nu'_w|_{C'_w \cap C_{w\widehat{w}w}}, \quad \widehat{w} \in \widehat{A}_w, \ \ 0 \le i \le |w| + |\widehat{w}| - 3. \tag{23}$$

It is clear that the measure ν_w^∞ is thus well-defined (as a measure on X) and T-invariant. Besides, replacing ν'_w by $\overline{\mu}'_w$ on the right hand side of (23) yields the measure $\overline{\mu}_w^\infty = \overline{\mu}/\overline{\mu}(C'_w)$ on the left hand side.

Compare $\overline{\mu}_w^\infty(C_w^\infty)$ and $\nu_w^\infty(C_w^\infty)$. Since $\overline{\mu}$ is an ergodic measure and C_w^∞ is a T-invariant set of positive measure, we have $\overline{\mu}(C_w^\infty) = 1$, i.e.,

$$\sum_{\widehat{w} \in \widehat{A}_w} \sum_{i=0}^{|w|+|\widehat{w}|-3} \overline{\mu}\left(T^i\left(C'_w \cap C_{w\widehat{w}w}\right)\right) = \sum_{\widehat{w} \in \widehat{A}_w} (|w| + |\widehat{w}| - 2)\overline{\mu}\left(C'_w \cap C_{w\widehat{w}w}\right) = 1.$$

Hence, by (21), (23), and the definition of $\overline{\mu}'_w$,

$$(1 - \varepsilon) \nu_w^\infty(C_w^\infty) \le \overline{\mu}_w^\infty(C_w^\infty) = \frac{1}{\overline{\mu}(C'_w)} \le (1 + \varepsilon) \nu_w^\infty(C_w^\infty). \tag{24}$$

We denote the normalized version of ν_w^∞ by $\overline{\nu}_w$, i.e., $\overline{\nu}_w = \nu_w^\infty/\nu_w^\infty(C_w^\infty)$.

Our next goal is to show that some finite-dimensional distributions of the T-invariant probability measures $\overline{\nu}_w$ and $\overline{\mu}$ are close to each other.

By the definition of $\overline{\nu}_w$ and ν_w^∞

$$\frac{\overline{\nu}_w(T^i(C_w' \cap C_{w\widehat{w}w}))}{\overline{\mu}(T^i(C_w' \cap C_{w\widehat{w}w}))} = \frac{\overline{\nu}_w(C_w' \cap C_{w\widehat{w}w})}{\overline{\mu}(C_w' \cap C_{w\widehat{w}w})} = \frac{\nu_w'(C_w' \cap C_{w\widehat{w}w})\overline{\mu}(C_w')}{\nu_w^\infty(C_w^\infty)\overline{\mu}(C_w')\overline{\mu}(C_w' \cap C_{w\widehat{w}w})}.$$

Since $\overline{\mu}$ is an ergodic measure, one can everywhere replace $\overline{\mu}(C_w')$ by $\overline{\mu}(C_w)$ and $\overline{\mu}(C_w' \cap C_{w\widehat{w}w})$ by $\overline{\mu}(C_{w\widehat{w}w})$. From this, using (21) and (24), we obtain

$$\frac{1-\varepsilon}{1+\varepsilon} \leq \frac{\overline{\mu}(T^i(C_w' \cap C_{w\widehat{w}w}))}{\overline{\nu}_w(T^i(C_w' \cap C_{w\widehat{w}w}))} \leq \frac{1+\varepsilon}{1-\varepsilon}, \qquad \widehat{w} \in \widehat{A}_w. \tag{25}$$

Notice that $\overline{\nu}_w$ is also an ergodic measure. Indeed, by Proposition 4.1 the dynamical system $(C_w', T_{C_w'}, \nu_w')$ is isomorphic to the Bernoulli shift (Y_w, σ, ν^w) and hence ergodic. But $(X, T, \overline{\nu}_w)$ is by construction an integral automorphism over $(C_w', T_{C_w'}, \nu_w')$, which is known to be ergodic together with the automorphism in the base. Therefore (25) can be rewritten in the form

$$\frac{1-\varepsilon}{1+\varepsilon} \leq \frac{\overline{\mu}(T^i C_{w\widehat{w}w})}{\overline{\nu}_w(T^i C_{w\widehat{w}w})} \leq \frac{1+\varepsilon}{1-\varepsilon}, \qquad \widehat{w} \in \widehat{A}_w. \tag{26}$$

Consider an arbitrary word $\widetilde{w}$ in G and denote the set of pairs $(\widehat{w}, i)$ such that

$$\widehat{w} \in \widehat{A}_w, \qquad 0 \leq i \leq |w| + |\widehat{w}| - 3, \qquad C_{\widetilde{w}} \cap T^i C_{w\widehat{w}w} \neq \emptyset$$

by $\widehat{A}_{w,\widetilde{w}}$. Since $\overline{\mu}(C_w^\infty) = 1$, $\overline{\mu}$-almost every point x of the cylinder $C_{\widetilde{w}}$ belongs to C_w^∞. For such x, there exists a unique pair $(\widehat{w}, i) \in \widehat{A}_{w,\widetilde{w}}$ such that $x \in T^i C_{w\widehat{w}w}$. Thus the cylinders $C_{\widetilde{w}} \cap T^i C_{w\widehat{w}w}$ with $(\widehat{w}, i) \in \widehat{A}_{w,\widetilde{w}}$ constitute a $\overline{\mu}$-mod 0 partition of $C_{\widetilde{w}}$. Assume that $|\widetilde{w}| \leq |w|$. Then either $C_{\widetilde{w}} \cap T^i C_{w\widehat{w}w} = \emptyset$, or $T^i C_{w\widehat{w}w} \subset C_{\widetilde{w}}$. Hence

$$\overline{\mu}(C_{\widetilde{w}}) = \sum_{(\widehat{w},i)\in\widehat{A}_{w,\widetilde{w}}} \overline{\mu}\big(C_{\widetilde{w}} \cap T^i C_{w\widehat{w}w}\big).$$

For the same reasons a similar formula holds for the measure $\overline{\nu}_w$ in place of $\overline{\mu}$. Moreover, in both sums on the right hand side the same terms equal zero, while the non-zero terms equal $\overline{\mu}(C_{w\widehat{w}w})$ and $\overline{\nu}_w(C_{w\widehat{w}w})$, respectively. Hence by (26), for every $\widetilde{w}$ with $|\widetilde{w}| \leq |w|$,

$$1 - 2\varepsilon \leq \frac{1-\varepsilon}{1+\varepsilon} \leq \frac{\overline{\mu}(C_{\widetilde{w}})}{\overline{\nu}_w(C_{\widetilde{w}})} \leq \frac{1+\varepsilon}{1-\varepsilon} \leq 1 + 3\varepsilon, \tag{27}$$

where the last inequality holds when $\varepsilon \leq 1/3$.

4. Now we are going to compare the integrals $\overline{\nu}_w(f)$ and $\overline{\mu}(f)$. Using properties of f, one can find a number $n(\varepsilon) \in \mathbb{N}$ and a function f_ε such that $f_\varepsilon(x)$ depends on $x_0, \ldots, x_{n(\varepsilon)}$ only and $|f(x) - f_\varepsilon(x)| \leq \varepsilon$ for all $x \in X$. Then

$$
\begin{aligned}
|\overline{\mu}(f) - \overline{v}_w(f)| &= \left| \overline{\mu}(f) - \overline{\mu}(f_\varepsilon) + \overline{\mu}(f_\varepsilon) \right. \\
&\quad \left. - \overline{v}_w(f_\varepsilon) + \overline{v}_w(f_\varepsilon) - \overline{v}_w(f) \right| \\
&\leq \overline{\mu}(|f - f_\varepsilon|) + \overline{v}_w(|f - f_\varepsilon|) + |\overline{\mu}(f_\varepsilon) - \overline{v}_w(f_\varepsilon)| \\
&\leq 2\varepsilon + |\overline{\mu}(f_\varepsilon) - \overline{v}_w(f_\varepsilon)|.
\end{aligned}
\tag{28}
$$

The function f_ε is chosen so that it is constant on every cylinder of the form $C_{\widetilde{w}}$, where $|\widetilde{w}| = n(\varepsilon)$. We number these cylinders: $C_{\widetilde{w}_1}, C_{\widetilde{w}_2}, \ldots$, and denote the value of f_ε on $C_{\widetilde{w}_i}$ by $f_{\varepsilon,i}, i = 1, 2, \ldots$ Now, by (27), (28), and the definition of f_ε,

$$
\begin{aligned}
|\overline{\mu}(f) - \overline{v}_w(f)| &\leq 2\varepsilon + \sum_i \left| \overline{\mu}(C_{\widetilde{w}_i}) - \overline{v}_w(C_{\widetilde{w}_i}) \right| f_{\varepsilon,i} \\
&\leq 2\varepsilon + 3\varepsilon \sum_i \overline{\mu}(C_{\widetilde{w}_i}) f_{\varepsilon,i} \\
&\leq 3\varepsilon \left(1 + \overline{\mu}(f_\varepsilon)\right) \leq 3\varepsilon \left(1 + \overline{\mu}(f) + \varepsilon\right),
\end{aligned}
$$

which can be made arbitrary small by choosing a small ε.

By the definition of the measure v^w it is a maximal measure for the suspension flow (σ, ψ^w), i.e.,

$$
\frac{h_{v^w}(\sigma)}{v^w(\psi^w)} = h_{\text{top}}(\sigma, \psi^w) = h_{\text{top}}(\sigma, \varphi^w).
$$

(For the last equality see (17).) Using (16), we obtain

$$
\left| \frac{h_{v^w}(\sigma)}{v^w(\varphi^w)} - \frac{h_{v^w}(\sigma)}{v^w(\psi^w)} \right| \leq \frac{\varepsilon h_{v^w}(\sigma)}{v^w(\varphi^w) v^w(\psi^w)} \leq \frac{2\varepsilon}{c} h_{\text{top}}(\sigma, \varphi^w).
$$

(For the definition of ψ^w see step 3 of this proof.) Hence

$$
\left| \frac{h_{v^w}(\sigma)}{v^w(\varphi^w)} - h_{\text{top}}(\sigma, \varphi^w) \right| \leq \frac{2\varepsilon}{c} h_{\text{top}}(\sigma, \varphi^w).
\tag{29}
$$

By Propositions 4.1 and 4.3 the suspension flows (σ, φ^w) and $(T_{C'_w}, f_w)$ are isomorphic. From this it follows that

$$
h_{\text{top}}(\sigma, \varphi^w) = h_{\text{top}}\left(T|_{C_{\widetilde{w}}^\infty}, f|_{C_{\widetilde{w}}^\infty}\right)
$$

and that the suspension flows with invariant measures (σ, φ^w, v^w) and $(T, f, \overline{v}_w)$ are isomorphic as well (see the definition of the measure $\overline{v}_w$). The latter fact together with (29) imply that

$$
\left| \frac{h_{\overline{v}_w}(T)}{\overline{v}_w(f)} - h_{\text{top}}(T, f) \right| \leq \frac{2\varepsilon}{c} h_{\text{top}}(T, f).
\tag{30}
$$

5. It remains to estimate $h_{\overline{\mu}}(T)$ from below.

We take an arbitrary sequence of positive numbers $\varepsilon(n) \leq 1/3, n \geq 1$, such that $\lim_{n \to \infty} \varepsilon(n) = 0$ and denote $w_n = w_{\varepsilon(n)}$ (see Corollary 4.7). Let also

$$r_n = |w_n|, \quad \alpha_n = 3\varepsilon(n), \quad \mu_0 = \overline{\mu}, \quad \mu_n = \overline{\nu}_{w_n}, \quad h = \overline{\mu}(f) h_{\text{top}}(T, f).$$

Without loss of generality we can assume that $|w_n| \to \infty$ as $n \to \infty$. Using (28), (27), and (30) with $\varepsilon = \varepsilon(n)$, it is easy to check that the assumptions of Lemma 3.4 are satisfied. By this lemma $h_{\mu_0}(T) \geq \overline{\mu}(f) h_{\text{top}}(T, f)$, so that

$$\frac{h_{\overline{\mu}}(T)}{\overline{\mu}(f)} \geq h_{\text{top}}(T, f).$$

But the opposite inequality holds for every measure from $\mathcal{I}_f(T)$. Thus the proof is completed.

References

1. Bufetov, A.I., Gurevich, B.M.: On the measure with maximum entropy for the Teichmüller flow on the moduli space of Abelian differentials. Funct. Anal. Appl. **42**(3), 224–226 (2008)
2. Bufetov, A.I., Gurevich, B.M.: Existence and uniqueness of the measure of maximal entropy for the Teichmüller flow on the moduli space of Abelian differentials. Sb. Math. **202**(7), 935–970 (2011)
3. Buzzi, J., Sarig, O.: Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps. Ergodic Theory Dynam. Systems **23**(5), 1383–1400 (2003)
4. Cornfeld, I.P., Fomin, S.V., Sinai, Ya.G.: Ergodic Theory. Grundlehren der Mathematischen Wissenschaften, vol. 245. Springer, Berlin (1982)
5. Gurevich, B.M.: A variational characterization of one-dimensional countable state Gibbs random fields. Z. Wahrsch. Verw. Gebiete **68**(2), 205–242 (1984)
6. Gurevich, B.M.: Weak approximation of an invariant measure and a low boundary of the entropy (2014). arXiv:1408.1609
7. Gurevich, B., Katok, S.: Arithmetic coding and entropy for the positive geodesic flow on the modular surface. Mosc. Math. J. **1**(4), 569–582 (2001)
8. Katok, A., Hasselblatt, B.: Introduction to the Modern Theory of Dynamical Systems. Encyclopedia of Mathematics and its Applications, vol. 54. Cambridge University Press, Cambridge (1995)
9. Sarig, O.M.: Bernoulli equilibrium states for surface diffeomorphisms. J. Mod. Dyn. **5**(3), 593–608 (2011)
10. Sarig, O.M.: Symbolic dynamics for surface diffeomorphisms with positive entropy. J. Amer. Math. Soc. **26**(2), 341–426 (2013)
11. Savchenko, S.V.: Special flows constructed from countable topological Markov chains. Funct. Anal. Appl. **32**(1), 32–41 (1998)
12. Veech, W.A.: The Teichmüller geodesic flow. Ann. Math. **124**(3), 441–530 (1986)